Accepted Manuscript

CHIS: A big data infrastructure to manage digital cultural items

Aniello Castiglione, Francesco Colace, Vincenzo Moscato, Francesco Palmieri

PII: S0167-739X(17)30560-5
DOI: http://dx.doi.org/10.1016/j.future.2017.04.006
Reference: FUTURE 3408

To appear in: *Future Generation Computer Systems*

Received date: 8 January 2017
Revised date: 29 March 2017
Accepted date: 4 April 2017

Please cite this article as: A. Castiglione, F. Colace, V. Moscato, F. Palmieri, CHIS: A big data infrastructure to manage digital cultural items, *Future Generation Computer Systems* (2017), http://dx.doi.org/10.1016/j.future.2017.04.006

CHIS: A Big Data Infrastructure to Manage Digital Cultural Items

Aniello Castiglione[a,*], Francesco Colace[b], Vincenzo Moscato[c], Francesco Palmieri[a]

[a] *University of Salerno, Dipartimento di Informatica, Via Giovanni Paolo II, 132 - 84084 Fisciano (SA), Italy*
[b] *University of Salerno, Dipartimento di Ingegneria Industriale, Via Giovanni Paolo II, 132 - 84084 Fisciano (SA), Italy*
[c] *University of Naples, Dip. di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Via Claudio 21, - 80125 Napoli, Italy*

Abstract

In this paper, we describe CHIS (*Cultural Heritage Information System*), a big data infrastructure that can be used to query, browse, analyze and process digital contents related to cultural heritage from a set of heterogeneous and distributed repositories. CHIS is characterized by the following technical features: capability to gather information from distributed and heterogeneous data sources (e.g., Sensor Networks, Social Media Networks, Digital Libraries and Archives, Multimedia Collections, Web Data Services, etc.); advanced data management techniques and technologies; ability to provide useful and personalized data to users based on their preferences and context; advanced information retrieval facilities, data analytics and other utilities/services, according to the SOA paradigm. By means of a set of ad-hoc APIs, and value-added data processing and analytics services, our system can support several applications: mobile multimedia guides for cultural environments, web portals to promote the cultural heritage of a given organization, multimedia recommender and storytelling systems and so on. We discuss the main ideas that characterize the system, showing its use for several applications.

Keywords: Big Data, Cultural Heritage, Resource Management, Big Data Analytics, SOA, NoSQL

1. Introduction

Italy boasts one of the largest and most priceless cultural heritage in the world. This invaluable resource can be opportunely protected, preserved and promoted by "embedding" it in the digital ecosystem of a *Smart City*, where economic, tourist, recreational and logistic aspects have to be considered all together.

As highlighted by recent publications and reiterated by European projects discussing Smart Cities [1, 2, 3, 4], the adoption of *Future Internet* technologies, particularly the paradigms of the *Internet of Things* and *Internet of Services*, now represents the "de facto standard" in the design and implementation of IT platforms that can provide effective support to the "smartness" of a city. In such a context, it is possible to design *context-aware* services that take into account both the surrounding environment, whose state is captured by sensors, and the characteristics of the users. These services are then all accessible through a *Cloud Computing* environment [5]. In addition, according to the vision of *participatory sensing*, [6] mobile devices of the latest generations (e.g., smart-phones, tablets, etc.) form an interactive network that allows users to access, analyze and share information and knowledge with an "active" role. Eventually, following the recommendations of the W3C *Semantic Web* framework [7, 8], in order to allow the enormous amount of data collected (*Big Data*[9, 10]) in a smart environment to be used by different applications, the data must be properly processed and stored in the form of *linked open data*, in order to facilitate both access and semantic processing.

A number of proposals, which focus on how the discussed technological solutions should be applied to the cultural domain, has been already presented for the Italian heritage.

*Corresponding author: Aniello Castiglione, Department of Computer Science, University of Salerno, Via Giovanni Paolo II, 132 I-84084 Fisciano (SA), Italy - Phone: +39089969594, Fax: +39089969821, email: castiglione@acm.org, castiglione@ieee.org

Email addresses: castiglione@ieee.org (Aniello Castiglione), fcolace@unisa.it (Francesco Colace), vmoscato@unina.it (Vincenzo Moscato), fpalmieri@unisa.it (Francesco Palmieri)